Supplementary Material to Accompany: "The Role of Labor and Marriage Markets, Preference Heterogeneity and the Welfare System in the Life Cycle Decisions of Black, Hispanic and White Women

by

Michael P. Keane
ARC Federation Fellow, University of Technology Sydney
and
Kenneth I. Wolpin
University of Pennsylvania

April, 2007

Abstract

Using data from the NLSY79, we structurally estimate a dynamic model of the life cycle decisions of young women. The women make sequential joint decisions about school attendance, work, marriage, fertility and welfare participation. We use the model to perform counterfactual simulations designed to shed light on three questions: (1) How much of observed minority-majority differences in behavior can be attributed to differences in labor market opportunities, marriage market opportunities, and preference heterogeneity? (2) How does the welfare system interact with these factors to augment those differences? (3) How can new cohorts that grow up under the new welfare system (TANF) be expected to behave compared to older cohorts?

The authors are grateful for support from NICHD under grant HD-34019, from the ARC under grant FF0561843, and several grants from the Minnesota Supercomputer Institute. An early version of this work appeared under the title "Public Welfare and the Life Cycle Decisions of Young Women."

Keywords: female life-cycle behavior, labor market opportunities, marriage market opportunities, public welfare

JEL Codes: J1, J2, J3

Appendix SA: Welfare Benefit Rules

In order to estimate the benefit schedules given by equation (5) in the main text, and the evolutionary rules governing changes in benefit parameters given by equation (6), we collected information on the rules governing AFDC and Food Stamp eligibility and benefits in each of the 50 states for the period 1967-1990. We then simulated a large data set of hypothetical women, with different numbers of children, and different levels of labor and non-labor income, and calculated their welfare benefits according to the exact rules in each State and year.[1] We calculated the sum of monthly benefits from AFDC and Food Stamps, and expressed these monthly benefit amounts in 1987 New York equivalent dollars. The resulting simulated data was used to estimate the approximate benefit schedule given by (5) separately for each State and year. Thus, for each state, s, we obtain an estimate of the benefit rule parameters, $b_{t0}^{s}, b_{t1}^{s}, b_{t2}^{s}, b_{t3}^{s}, b_{t4}^{s}$, for each year t.[2] Given the estimates of the benefit rule parameters, we then estimated (6), the evolutionary rule (ER).

Table A.1 reports summary statistics of the benefit rule parameters $b_{t0}^{s}, b_{t1}^{s}, b_{t2}^{s}, b_{t3}^{s}, b_{t4}^{s}$ by State. We report the mean of each parameter over the 1967-1990 sample period, as well as the standard deviation and the minimum and maximum. Table A.2 reports our estimates of the evolutionary rule (i.e., the VAR for the 5 parameters) for each of the 5 States.

[1] Deductions for child care expenses and work expenses, as well as various other income disregards that existed under the AFDC program, were also factored into these calculations. EITC was also factored in, but this was quite trivial prior to the expansion in the 1993-94 period.

[2] The approximation given by (5) fits the monthly benefit data quite well, with R-squared statistics for the first line segment mostly above .99 and for the second, mostly about .95. These regressions are available on request.

Appendix SB: Estimation Method

The numerical solution to the agents' maximization problem provides (approximations to) the Emax functions that appear on the right hand side of (8). The alternative-specific value functions, V_a^j for j=1,..,J , which are sums of current payoffs and the discounted Emax functions, are known to the agents in the model. But the econometrician does not observe all the factors that enter the current payoff expressions U_a^j. In general, the econometrician does not observe the random preference shocks, the part- and full-time wage offer shocks, the earnings shock of the husband and the income shock of her parents. Whether particular alternatives are available depends on the implicit shocks governing whether a part- and/or full-time job offer is received, whether a marriage offer is received and whether a parental co-residence offer is received.

Thus, conditional on the deterministic part of the state space, the probability that an agent is observed to choose option k takes the form of an integral over the region of the several-dimensional error space such that k is the preferred option. The error space over which the econometrician must integrate depends on which option k the agent is observed to chose. For example, if the work option is chosen, then the wage offer is observed by us, and the wage shock is not in the subset over which we must integrate. In that case, the likelihood contribution for the observation also includes the density of the wage error. If the woman is married, then we observe the husband's income, we do not integrate over the husband's income shock, and the likelihood contribution includes the husband's income density.

As noted, the choice set contains as many as 36 elements, but our model imposes a factor structure where a much smaller number of errors (i.e., the wage shocks and the 5-element vector of preference shocks over leisure, school, marriage, fertility and welfare) determines choices.[3] It is well known that evaluation of choice probabilities is computationally burdensome when the number of alternatives is large. Recently, highly efficient smooth unbiased probability simulators, such as the GHK method (see, e.g., Keane (1993, 1994)), have been developed for these situations. Unfortunately, the GHK method, as well as other smooth unbiased simulators, rely on a structure in which each of the J-1 mutually exclusive alternatives have a value that is a

[3] Note that, despite this factor structure, the likelihood is not "degenerate" (i.e., meaning that no feasible choice has zero probability). This is because each of the five discrete choices (work, school, fertility, marriage welfare participation) has an associated error term whose distribution (i.e., Normal) covers the real line. Thus, there always exists a configuration of the errors that can rationalize any 5-element vector of choices that might be observed in the data.

strictly monotonic function of a single stochastic term, and that the (J-1)•(J-1) variance-covariance matrix of the error terms have full rank. This is not true here, because the alternatives have values that cannot be written as a strictly monotonic function of a single error.[4]

Furthermore, as discussed in Keane and Moffitt (1998), in estimation problems where the number of choices exceeds the number of error terms, the boundaries of the region of integration needed to evaluate a particular choice probability are generally intractably complex. Thus, given our model, the most practical method to simulate the probabilities of the observed choice set would be to use a kernel smoothed frequency simulator. These were proposed in McFadden (1989), and have been successfully applied to models with large choice sets in Keane and Moffitt (1998), Erdem (1996), Keane and Wolpin (1997) and Eckstein and Wolpin (1999).[5]

However, in the present context, this approach is not feasible because of severe problems created by unobserved state variables. As noted, we do not always have complete histories of employment, schooling or welfare receipt for most of the cohorts back to age 14. Hence, the state variables of work experience, completed schooling and lagged welfare participation cannot always be constructed. In addition, parental co-residence and marital status are observed only once a year (every other period).

Further complicating the estimation problem, as we noted earlier, is that the youth's initial schooling level at age 14 is observed only for one of the 16 birth cohorts. It is well known that unobserved initial conditions, and unobserved state variables more generally, pose formidable problems for the estimation of dynamic discrete choice models (Heckman (1981)). If some or all elements of the state space are unobserved, then to construct conditional choice probabilities one must integrate over the distribution of the unobserved elements. Even in much simpler dynamic models than ours, such distributions are typically intractably complex.

In a previous paper (Keane and Wolpin (2001)), we developed a simulation algorithm that deals in a practical way with the problem of unobserved state variables. The algorithm is based on simulation of complete (age 14 to the terminal age) outcome histories for a set of artificial agents. An outcome history consists of the initial schooling level of the youth, S_0,

[4] One of us has written incorrectly elsewhere (see, e.g., Keane and Moffitt (1998)) that the GHK algorithm is only applicable when each alternative has a single error that is additive and the error covariance matrix is at least of rank J-1. A prime example is the multinomial probit model. But, additivity is in fact a much stronger condition than is required. Strict monotonicity is sufficient.

[5] Kernel smoothed frequency simulators are, of course, biased for positive values of the smoothing parameter, and consistency requires letting the smoothing parameter approach zero as sample size increases.

parental schooling, $\mathbf{S}^{\mathbf{z}}$, along with simulated values in all subsequent periods for all of the outcome variables in the model (school attendance, part- or full-time work, marriage, pregnancy, welfare participation, the woman's wage offer, the husband's earnings, both permanent and transitory components, parental co-residence and income). The construction of an outcome history can be described compactly as follows:

At the current trial parameter value, we simulate histories as follows:

1) Draw the youth's "type," which includes both her skill endowment and 5-element vector of preference parameters, as well as her initial schooling and parent's schooling, from a joint distribution;[6]

2) Draw the relevant set of random shocks necessary to compute the alternative-specific value functions at age a=1;

3) Choose the alternative with the highest alternative-specific value function;

4) Update the state variables based on the choice in (3);

5) Repeat steps (2) – (4) for a=2, ... , A;

We repeat steps (1) - (5) N times to obtain simulated outcome histories for N artificial persons. Denote by $\tilde{\mathbf{O}}^{\mathbf{n}}$ the simulated outcome history for the n[th] such person, so $\tilde{\mathbf{O}}^{\mathbf{n}} = (\Omega_0^{\mathbf{n}}, \tilde{\mathbf{O}}^{\mathbf{n}}_{\mathbf{a}=1}, ..., \tilde{\mathbf{O}}^{\mathbf{n}}_{\mathbf{a}=\mathbf{A}})$, for n = 1,..., N.

In order to motivate the estimation algorithm, it is useful to ignore for now the complication that some of the outcomes are continuous variables and that there are observed initial conditions and unobserved types. Let $\mathbf{O}^{\mathbf{i}}$ denote the observed outcome history for person i, which may include missing elements. Then, an unbiased frequency simulator of the probability of the observed outcome history for person i, $\mathbf{P}(\mathbf{O}^{\mathbf{i}})$, is just the fraction of the N simulated histories that are consistent with $\mathbf{O}^{\mathbf{i}}$. In this construction, missing elements of $\mathbf{O}^{\mathbf{i}}$ are counted as consistent with any entry in the corresponding element of $\tilde{\mathbf{O}}^{\mathbf{n}}$. Note that the construction of this simulator relies only on unconditional simulations. It does not require evaluation of choice probabilities conditional on state variables. Thus, unobserved state variables do not create a problem for this procedure.

[6] We do not draw from the "correct" joint distribution. Instead, we draw from an incorrect "source" distribution and adjust the draws so obtained using importance sampling weights. The virtue of this procedure, similar to Keane (1993, 1994), will become apparent below. The key support condition for importance sampling is that the source distribution put positive mass on each possible type/initial school/parent's school combination, which is easy to verify in this case.

Unfortunately, this algorithm is not practical. Because the number of possible outcome histories is huge, consistency of a simulated history with an actual history is an extremely low probability event. Hence, simulated probabilities will typically be 0, as will be the simulated likelihood, unless an impractically large simulation size is used (see Lerman and Manski 1981). In addition, the method breaks down completely if any outcome is continuous (e.g., the woman's wage offer), regardless of simulation size, because agreement of observed with simulated wages is a measure zero event.

We solve this problem by assuming, as seems apt, that all observed quantities are measured with error. With measurement error there is a nonzero probability that any observed outcome history might be generated by any simulated outcome history. Denote by $P(O^i|\tilde{O}^n)$ the probability that observed outcome history O^i is generated by simulated outcome history $\tilde{O}^n$. Then $P(O^i|\tilde{O}^n)$ is the product of classification error rates on discrete outcomes (and measurement error densities for the continuous variables) that are needed to make O^i and $\tilde{O}^n$ consistent. Observe that $P(O^i|\tilde{O}^n)>0$ for any $\tilde{O}^n$, given suitable choice of error processes. The specific measurement error processes that we assume are described below. The key point here is that $P(O^i|\tilde{O}^n)$ does not depend on the state variables at any age a, but only depends on the outcomes.

Using N simulated outcome histories we obtain the unbiased simulator:

$$\text{(B1)} \quad \hat{P}_N(O^i) = \frac{1}{N}\sum_{n=1}^{N} P(O^i|\tilde{O}^n).$$

Note that this simulator is analogous to a kernel-smoothed frequency simulator, in that $I(O^i=\tilde{O}^n)$ is replaced with an object that is strictly positive, but that is greater if $\tilde{O}^n$ is "closer" to O^i. However, the simulator in (B1) is unbiased because the measurement error is assumed to be present in the true statistical model.

It is straightforward to extend the estimation method to allow for unobserved heterogeneity. Assume that there are K types of women who differ in their permanent preferences for leisure, school, marriage, becoming pregnant and receiving welfare, as well as in their human capital "endowment" at age 14.[7] In addition, women also differ in terms of their initial schooling

[7] At a point in time, married women also differ in terms of the permanent unobservable component of their husband's human capital, μ^m in (4), which is fixed for the duration of a marriage. But this is not part of a woman's initial condition.

(taking on 4 values) and parental schooling (taking on14 values); initial schooling, as we have noted, is often unobserved. Thus, there are a total of 56·K possible initial conditions in simulation step (1) of the algorithm to generate histories. Let k = 1,..., 56·K index these initial conditions, and define π_k as the probability a person has initial condition k given the joint distribution of unobserved type, initial schooling and parental schooling assumed in the model.[8] Also, define π_{k0} as the proportion of agents with initial condition k simulated in step 1, and let k(n) denote the initial condition that was drawn in step 1 when simulating history n. Finally, let $\tilde{O}_k^n$ denote that the n[th] outcome history, which is simulated under the assumption the agent has initial condition k. Then, we can form the unbiased simulator:

$$\text{(B2)}\quad \hat{P}_N(O^i) = \frac{1}{N}\sum_{n=1}^{N} P(O^i|\tilde{O}_k^n)\frac{\pi_{k(n)}}{\pi_{k(n),0}}.$$

Observe that in (B2), the conditional probabilities $P(O^i|\tilde{O}_k^n)$ are weighted by the ratio of the probability of agents with initial condition k according to the model, π_k, to the probability of agents with initial condition k in the simulation, π_{k0}. As we discuss in Appendix A, we construct the joint distribution of latent type, initial schooling, and parental schooling using (i) a multinomial logit (MNL) for initial schooling conditional on parents' schooling, in conjunction with (ii) a MNL for type conditional on parent and initial schooling. Together, these logits generate the π_k.

It is important for probability simulators to be smooth functions of model parameters for several reasons.[9] The simulator in (B2) is a smooth function of the MNL parameters that determine the type proportions π_k, and this is a key virtue of using importance sampling in Step 1 of the algorithm for constructing histories.[10] Unfortunately, (B2) is not a smooth function of the structural parameters that determine choice probabilities conditional on initial conditions. This is because $P(O^i|\tilde{O}^n)$ will "jump" at points where a change in a model parameter causes the simulated outcome history $\tilde{O}^n$ to change discretely. However, this simulator can be made smooth in these parameters by applying a second importance sampling procedure. The idea is to hold the

[8] Parental schooling and initial schooling are assumed to be exogenous conditional on type.

[9] As discussed in McFadden (1989) and Keane (1994), smoothness allows construction of derivatives, which both speeds the search for an optimum and permits calculation of numerical standard errors. It also typically leads to more efficient simulators, and avoids problems created by zero simulated probabilities (see Lerman and Manski (1981)).

[10] Keane and Wolpin (1997, 2001) adopted the same approach to handling latent types.

simulated outcome histories fixed as the model parameters are varied, but to reweight them in an appropriate way.

Given an initial parameter vector θ_0 and an updated vector θ', the appropriate weight to apply to sequence $\tilde{O}^n$ is the ratio of the likelihood of simulated history n under θ' to that under θ_0. Such weights have the form of importance sampling weights (i.e., the ratios of densities under the target and source distributions), and are smooth functions of the model parameters. Further, it is straightforward to simulate the likelihood of an artificial history $\tilde{O}^n$ using conventional methods because the state vector is fully observed at all points along the history. The choice probabilities along a path $\tilde{O}^n$ are simulated using a kernel smoothed frequence simulator. As this construction renders $P(O^i|\tilde{O}^n)$ a smooth function of the model parameters, standard errors can be obtained using the BHHH algorithm.[11]

Lastly, it is necessary to describe our specific assumptions for the measurement error processes. First, we assume that discrete outcomes are subject to classification error. The structure we adopt is simply that there is some probability that the reported response category is the truth and some probability that it is not.[12] Second, we assume that the continuous variables are also subject to normally distributed measurement error. In particular, we assume that these errors are additive in the woman's log wage offer equation and in the husband's log income equation, while we assume that the parental income error is additive in levels. All measurement/classification errors are assumed to be serially independent and independent of each other. More details are given in Appendix A of the main text.

Finally, we discuss one subtle issue that arises in forming the simulated likelihood function using our algorithm. First, suppose that, according to simulated choice history $\tilde{O}^n$, a person's true choice at age a was not working, or not married, or not living with parents. Yet, in

[11] Despite the smoothness of the simulated likelihood function, estimation of the model proved difficult, as it was common for the search algorithm to become "hung up" on local maxima. We thus alternated between BHHH, the simplex algorithm, and simply moving sets of parameters by hand, switching methods whenever it appeared that one method had gotten "hung up." This laborious process ended when were no longer able to find any further improvement in the likelihood using any method. At this point the in-sample fit of the model also appeared to be quite reasonable, in the sense of capturing well many key features of the data. Our companion paper Keane and Wolpin (forthcoming) provides much more detail on model fit.

[12] To ensure that the measurement error is unbiased, the probability that the reported value is the true value must be a linear function of the predicted sample proportion (see Appendix A for details). Obviously, measurement error cannot be distinguished from the other model parameters in a non-parametric setting. As in the model without measurement error, identification relies on a combination of functional form and distributional assumptions, and exclusionary restrictions. Keane and Sauer (2005) have applied this algorithm successfully with more general classification error processes

the data, O^i we observed that the person is working, or is married, or is living with parents. Our method reconciles the two via classification error, and, for the discrete outcomes, the appropriate likelihood contribution is trivial: it is simply the probability the person is observed to work, be married or be living with parents, when in truth they are not. This probability is simply a function of the classification error rates constructed above.

But a more subtle problem arises in a case where the simulated history says a person was not working, or not married, or not living with parents, and, in the data, we not only observe a different discrete outcome, but also observe a wage, or husband earnings or parent's income. What is the density of an observed wage conditional on the person not actually working? Here, we make the simple assumption that such "falsely reported" continuous outcomes are drawn from the same distribution as that which governs the "true" continuous outcomes, except for a mean shift parameter that we estimate. We denote these mean shift parameters κ-w, κ-m, and κ-z for the woman's offer wage function, husband earnings function and parent's income function respectively. During estimation, κ-w never departed to any significant extent from zero, so we eventually pegged it at zero and report only κ-m, and κ-z in Appendix Table A of the main text.

References

Eckstein, Zvi and Kenneth I. Wolpin, "Why Youth Drop Out of High School: The Impact of Preferences, Opportunities, and Abilities," Econometrica, 67, Nov. 1999, 1295-1340.

Erdem, Tülin. "A Dynamic Analysis of Market Structure based on Panel Data," Marketing Science, 15, 1996, 359-378.

Heckman, James J. "The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process and Some Monte Carlo Evidence." in C. F. Manski and D. McFadden eds., Structural Analysis of Discrete Data with Econometric Applications, Cambridge, MA: MIT Press, 1981, 179-97.

Keane, Michael P. "Simulation Estimation for Panel Data Models with Limited Dependent Variables." in G.S. Maddala, C.R. Rao and H.D. Vinod eds., Handbook of Statistics 11, Amsterdam: Elsevier Science Publishers, 1993, 545-572.

___________. "A Computationally Practical Simulation Estimator for Panel Data." Econometrica, 62, January, 1994, 95-116.

Keane, Michael P. and Robert Moffitt. "A Structural Model of Multiple Welfare Program Participation and Labor Supply." International Economic Review, 39, August 1998, 553-590.

Keane, Michael P. and Robert Sauer. "A Computationally Practical Simulation Estimation Algorithm for Dynamic Panel Data Models With Unobserved Endogenous State Variables" Mimeo, Yale University, 2005.

Keane, Michael P. and Kenneth I. Wolpin, "The Career Decisions of Young Men." Journal of Political Economy, 105, June 1997, 473-522.

___________. "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment." International Economic Review, 42, November 2001, 1051-1103.

Lerman Steven R. and Charles F. Manski. "On the Use of Simulated Feequencies to Approximate Choice Probabilities." in C. F. Manski and D. McFadden eds., Structural Analysis of Discrete Data with Econometric Applications, Cambridge, MA: MIT Press, 1981, 305-19.

McFadden, Daniel. "A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration." Econometrica, 57, September, 1989, 995-1026.

Table A.1
Summary Statistics of Parameters of Benefits Rules by State: 1967-1990 (a,b)

		b_0	b_1	b_2	b_3	b_4
CA						
	μ	454	134	503	.64	166
	σ	53	9	47	.15	12
	Min	332	108	393	.24	143
	Max	517	148	579	.89	286
MI						
	μ	498	155	553	.63	193
	σ	78	16	118	.11	19
	Min	389	130	391	.53	146
	Max	649	181	744	.92	221
NY						
	μ	430	144	472	.63	179
	σ	38	24	65	.13	32
	Min	374	117	384	.48	142
	Max	522	182	590	.92	234
NC						
	μ	393	86	423	.52	110
	σ	42	18	83	.11	20
	Min	332	48	295	.41	84
	Max	462	111	545	.82	148
OH						
	μ	371	118	415	.58	143
	σ	26	12	71	.10	23
	Min	337	100	308	.47	114
	Max	415	143	539	.88	183

a. 1987 NY dollars
b. Based on Monthly AFDC plus Food Stamp Benefits

Table A.2
Evolutionary Rules for Benefit Parameters[a]

	CA					MI				
	b_{0t}	b_{1t}	b_{2t}	b_{3t}	b_{4t}	b_{0t}	b_{1t}	b_{2t}	b_{3t}	b_{4t}
$b_{0,t-1}$	.834 (.104)	.051 (.032)	-	-.00039 (.0006)	-	-.120 (.280)	-.086 (.050)	-.547 (.286)	-	-
$b_{1,t-1}$	.840 (.590)	.227 (.185)	-	-.00047 (.0034)	-	.446 (.903)	.774 (.164)	-.524 (.924)	-	-
$b_{2,t-1}$	-.322 (.130)	.041 (.040)	.640 (.128)	-.00040 (.0007)	-	.514 (.203)	.078 (.036)	1.04 (.207)	-	-
$b_{3,t-1}$	59.4 (19.4)	9.52 (6.12)	-	.673 (.114)	-	166.9 (67.6)	27.4 (12.3)	60.5 (69.1)	.614 (.117)	-
$b_{4,t-1}$	.496 (.404)	-.236 (.133)	-	.00601 (.002)	.469 (.152)	.468 (.870)	-.070 (.163)	1.71 (.896)	-	.800 (.101)
Constant	83.3 (55.3)	105.5 (18.4)	178.7 (64.8)	-.749 (.317)	87.6 25.4)	216.2 (124.8)	65.6 (23.9)	28.6 (129.3)	-.233 (.075)	38.1 (19.6)
R^2	.88	.53	.48	.60	.23	.89	.84	.94	.50	.74
P. Value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
Mean	454	134	503	.64	166	498	155	553	.63	193
RMSE	17.1	5.9	33.5	.087	10.3	25.9	6.2	28.5	.065	10.0

Table A.2, continued

	NY					NC				
	b_{0t}	b_{1t}	b_{2t}	b_{3t}	b_{4t}	b_{0t}	b_{1t}	b_{2t}	b_{3t}	b_{4t}
$b_{0,t-1}$	.851 (.065)	-	-	-	-	1.72 (.134)	.236 (.064)	2.18 (.328)	-.00249 (.0007)	.533 (.137)
$b_{1,t-1}$	-	.891 (.031)	-	-	-	-2.59 (.449)	.267 (.216)	-5.85 (1.10)	.00230 (.0026)	-.829 (.462)
$b_{2,t-1}$	-	-	.856 (.072)	-	-	-.446 (.090)	-.079 (.043)	-.619 (.221)	.00090 (.0005)	-.203 (.092)
$b_{3,t-1}$	-	-	-	.665 (.105)	-	201.0 (25.6)	77.3 (12.3)	144.1 (62.9)	.360 (.149)	86.7 (26.4)
$b_{4,t-1}$	-	-	-	-	.860 (.041)	1.38 (.381)	.287 (.183)	3.27 (.934)	-.00055 (.002)	1.07 (.392)
Constant	64.7 (28.6)	13.1 (4.70)	63.3 (35.2)	-.202 (.068)	22.1 (7.75)	77.1 (27.1)	14.1 (13.1)	37.1 66.6)	.141 (.158)	-14.3 (27.9)
R^2	.61	.92	.73	.54	.91	.97	.95	.95	.75	.86
P. Value	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00
Mean	430	144	472	.63	179	393	86	423	.52	110
RMSE	22.9	6.4	33.3	.074	8.7	7.3	3.5	17.8	.042	7.5

Table A.2, continued

	OH				
	b_{0t}	b_{1t}	b_{2t}	b_{3t}	B_{4t}
$b_{0,t-1}$	-.623 (.218)	.019 (.069)	-.045 (.312)	-	-
$b_{1,t-1}$	-.242 (.805)	.539 (.256)	-2.79 (1.15)	-	-
$b_{2,t-1}$	-.022 (.168)	-.027 (.053)	.126 (.241)	-	-
$b_{3,t-1}$	5.02 (32.3)	23.5 (10.3)	-144.6 (46.2)	.552 (.116)	-
$b_{4,t-1}$	1.19 (.560)	.230 (.181)	2.93 (.801)	-	.904 (.082)
Constant	261.8 (49.7)	38.9 (16.6)	195.6 (71.0)	-.243 (.069)	12.5 (12.0)
R^2	.79	.75	.94	.48	.84
P. Value	0.00	0.00	0.00	0.00	0.00
Mean	371	118	415	.58	143
RMSE	11.4	5.7	16.0	.056	9.0